

Centre for Policy Research on
Men and Boys

**International
Literature Review
on Boys' Education:
Readers' Guide**

Jordan Hill

Centre for Policy Research on Men and Boys

The Centre for Policy Research on Men and Boys (CPRMB) is a research organisation and think tank dedicated to understanding and addressing issues that uniquely affect men and boys in the UK. Its vision is a world where men and boys of all backgrounds thrive in their families and communities, one where the sexes can rise together by supporting each other.

It will do this by supporting the development of new research, thinking and insight into policy affecting men and boys. This is by focusing on critical areas where they face unique challenges, promoting policy change, and fostering public awareness and understanding. These are in seven core areas: Economy, Employment and Skills; Education; Health; Fatherhood and Family; Criminal Justice; Male Identity; and Portrayal of Men in Media and Culture.

This report was authored by Jordan Hill. Jordan is a Research Fellow at the Centre, specialising in education policy. A former Education Policy Analyst at the OECD Directorate for Education and Skills, he has also held research and policy roles in Brussels and London.

Reader's Guide

This literature review is a comprehensive, cross-national synthesis of evidence explicitly focused on boys'¹ educational outcomes and the factors that shape them in high-income countries². It was commissioned by the Centre for Policy Research on Men and Boys (CPRMB) to address a significant gap in the policy research landscape:

Boys in high-income countries are more likely than girls to fall behind in school. This is clearest in reading, but it shows up in other areas, including dropout rates and further study. We have known for many decades that these gender gaps exist but do not fully understand why. Factors like poverty, gender norms, and the school environment all play a role – but no framework exists that maps how they interact. Girls' education has rightly received significant policy attention, and that work must continue. Understanding boys' underperformance is not a competing priority. It is part of the same commitment to every child having a fair chance. This literature review builds a framework to give policymakers a stronger evidence base for targeting their responses to the challenge of boys' education outcomes.

The review draws on peer-reviewed research, national assessment data and international large-scale student assessments across high-income economies spanning North America, Europe, Asia-Pacific and Latin America³. It is structured as three linked reports, published iteratively:

- **Part I** maps the scale and nature of gender gaps across cognitive, non-cognitive and attainment outcomes in high-income countries, establishing what we know about boys' educational outcomes relative to girls.
- **Part II** examines the contextual factors that shape these outcomes, including social norms, education system design, socioeconomic factors and developmental patterns, and the policy levers that can influence them.
- **Part III** investigates the drivers of student learning closest to boys' day-to-day experience of schooling, including the classroom environment, family and home, wellbeing and engagement, as well as what policy can do to influence them.

Gender-specific policy support is not without precedent. Across high-income countries, significant investments have already been made over many years to try to level the playing field and enable more girls to reach their full potential. Sustained investments have sought to expand girls' participation and achievement, particularly in STEM, through curricular reform, extracurricular programmes and confidence-building initiatives (Hammond et al., 2020). Although there is much work still to do, these efforts have coincided with positive changes. In the 1980s, data from 15 OECD countries shows that 15% of women aged 25-64 years had a tertiary education, compared to

¹ Unless otherwise stated, this review uses the terms 'boys', 'men' and 'males' to refer to individuals assigned male at birth who continue to identify as male.

² Every year, the World Bank Group classifies the world's economies into four income groups: low, lower-middle, upper-middle, and high. These classifications, updated each year on July 1, are based on the previous year's Gross National Income (GNI) per capita, expressed in U.S. dollars. For the 2026 fiscal year, high-income economies are those with more than a GNI per capita of \$13,935.

³ Primarily countries participating in Organisation for Economic Co-operation and development (OECD) and International Association for the Evaluation of Educational Achievement (IEA) international large-scale assessments and surveys. Countries that participated in the individual rounds of each assessment can be found on the IEA and OECD websites at <https://www.iea.nl/studies> and <https://www.oecd.org/en/about/members-partners.html>

19% of men. By the 2020-23 period, the same set of 15 countries saw 46% of women with a tertiary education, compared to 39% of men (OECD, 2025a). Across a similar period, the median gender wage gap across OECD countries fell from around 19% in 1995 to 11% today, a reduction of just over 40% relative to its mid-1990s level (OECD, 2025b). These trends reflect a concerted effort to break down the educational barriers to women's success in life and work. Doing so brings huge benefits to individuals, societies and economies.

Just as education research resists individualising responsibility for girls' academic success, the same critical lens must be applied to boys. Researchers and educators have rightly cautioned against implicitly blaming "girls for being girls," including claims that girls are "overly compliant", "excessively self-critical", or driven by an "innate desire to please others" (Hattie & Anderman, 2019). When confronted with systematic male underperformance across countries and decades, it is insufficient to attribute these outcomes to individual failures. Instead, the overrepresentation of boys among low performers should be understood as a systemic and global challenge, shaped by broader structures of inequality, masculinities, schooling cultures, and societal expectations.

There are important economic and social knock-on effects of ignoring systematic male underperformance in education. Boys with stronger literacy skills tend to have wider educational, career and life opportunities as well as better health outcomes and longer life expectancy (Brozo, 2019; Gilbert et al., 2018; Stenberg, 2025). It is estimated that, if boys made the same educational progress as girls while at school, their long-term annual productivity would be 1.3 percent higher; maintained over the course of a decade, this represents an increase in total production of 13.9 percent (Gregory & Welmond, 2021).

Conceptual framework

The conceptual framework used in this literature review (Figure 1) sees gender as a socially embedded characteristic that operates through norms, expectations and institutional practices, shaping education outcomes. This conceptualisation is crucial for three main reasons.

Firstly, gender alone is often not a particularly strong predictor of student outcomes but it can compound the risk of educational underperformance for students exposed to multiple forms of disadvantage. For example, gender does not exert anywhere near the same degree of influence on educational outcomes as socio-economic status (SES). Looking at GCSE outcomes in England, it is estimated that well over half of the variation in grades is accounted for by differences in SES, while gender accounts for around a fifth of the variation (Farquharson et al., 2024). However, the magnitude of impact that SES has on outcomes does not mean other factors such as gender, race, migration background, disability status, language spoken at home and place-based conditions deserve less attention. In fact, the impact SES has on education outcomes is heavily mediated by such factors (Gregory & Welmond, 2021; Jha et al., 2017). Students with similar socio-economic profiles can nonetheless face very different educational trajectories. Taking an intersectional approach is crucial to understand why certain groups of boys are likely to struggle at school, while others are not.

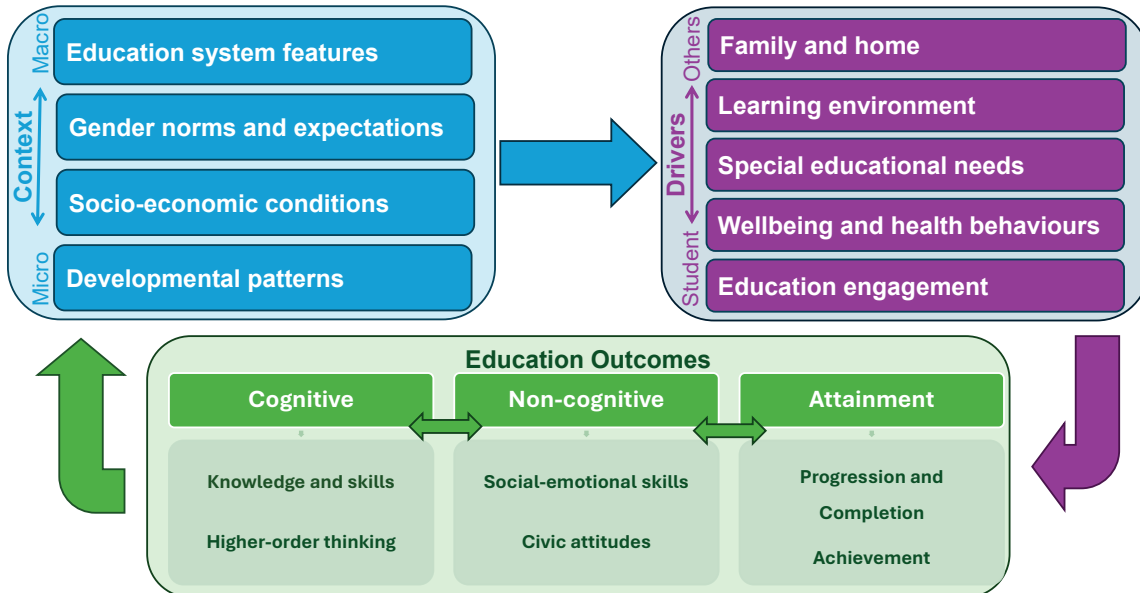
Secondly, although it is important to highlight the disproportionate number of boys who are struggling as students, evidence from psychology suggests that males and females are similar on most psychological variables (Hyde, 2014). Where differences are moderate-to-large in magnitude, they tend to relate to specific traits, such as sensation seeking and physical aggression (Hyde, 2014). In education, the prevailing and well-evidenced view is that boys and girls are more alike than they are different, and that there is substantially more variation within groups of boys and within groups of girls than there is between the two groups (Hattie, 2008). If gendered expectations are permitted to dominate judgments about ability, mathematically talented girls and reading-talented boys may be overlooked by parents and teachers alike (Hattie, 2008). Such categorisation would be detrimental to students of both genders while obscuring the “missing middle” of male students who neither excel nor fail but simply “get on with it” (Roberts, 2012). The model avoids seeing student gender as a deterministic characteristic, which would wrongfully lead to the implicit categorisation of all boys as poor students and all girls as good students.

Thirdly, since the core challenge of student underperformance concerns a minority of pupils, educational interventions must be adapted to local contexts if they are to be effective (Younger & Warrington, 2005). Legitimate concerns about the educational outcomes of some boys should not obscure the fact that many boys continue to do very well at school and boys are also often disproportionately represented among top-performing students (Autor et al., 2023; OECD, 2023b; Voyer & Voyer, 2014). Therefore, examining how risk profiles converge among groups of students is useful. Traits that are, on average, more prevalent among boys (such as higher levels of sensation seeking and physical aggression) may interact with socio-economic disadvantage, school climate, or local context in ways that compound the risk of educational underperformance. Taking this highly context-dependent view can help make policies and practices more tailored to student needs and less likely to have unforeseen and detrimental knock-on effects.

Integrating the above points, the model draws on ecological (Bronfenbrenner, 1981; UNICEF, 2022) and systems thinking approaches (Arnold & Wade, 2015; Richmond, 1994). These approaches are useful because they emphasise the complexity of interactions between actors (individuals and organisations), situated in social, institutional and cultural systems. The model is designed to move beyond average gender differences and instead identify how gendered risks accumulate and concentrate at the micro and macro levels among groups of students across the performance distribution, in particular poorer performers. It does this by using Context-Mechanism-Outcomes (Pawson & Tilley, 2010) and Context–Process–Outcomes (OECD, 2023a; Purves, 1987)

frameworks to situate boys' educational outcomes within a broader set of interacting contextual conditions, student-level drivers and education outcomes. The next section explains each part of the conceptual framework in more detail.

Figure 1 - Boys' Education Outcomes Framework



Contextual factors operate at both macro (e.g. societal) and micro (e.g. local, individual) levels and shape the conditions under which students develop and learn.

- **Gender norms and expectations** reflect the diverse array of socially constructed ideas about masculinity that influence behaviour, aspirations and responses to schooling. The model incorporates evidence on 'lad culture', norms around emotional expression and vulnerability, and the often-pragmatic approach boys take towards schooling.
- **Education system features** encompass the structural and institutional characteristics of schooling that shape learning pathways. These include assessment practices, stratification of students into different education pathways, school autonomy, curriculum frameworks, and initial teacher education.
- **Socio-economic conditions** capture both the school-level and place-based dimensions of disadvantage. This includes socio-economic disparities between regions and neighbourhoods.
- **Developmental patterns** refer to age-related and gender-linked trajectories in cognitive, emotional and behavioural development. These patterns describe average tendencies, rather than fixed outcomes.

The contextual factors above shape a set of **student-level drivers** that operate directly at the level of learning.

- **The learning environment** captures classroom and school-level experiences, including instructional quality, teacher beliefs, expectations and attitudes, and wider school organisation and climate. These factors shape how students engage with learning, how behaviour is interpreted by educators, and how difficulties are addressed.

- **Family and home conditions** include access to financial and cultural resources, parental expectations and support, the quality of school–parent relationships, and the experiences of at-risk children. These factors influence routines, support for learning and responses to emerging difficulties, often taking gendered pathways.
- **Special educational needs** are positioned as a driver in this framework to reflect the higher prevalence of learning and neurodevelopmental difficulties reported among boys. The model explicitly incorporates gendered patterns of identification and referral, school responses and subsequent educational trajectories. It does not engage with debates about over- or under-diagnosis, or with clinical literature on diagnostic criteria.
- **Wellbeing and health behaviours** include emotional wellbeing, diet, physical activity and sleep, time use and participation in extracurricular activities, and experiences of bullying and peer victimisation. These factors affect readiness to learn and engagement with schooling and often differ systematically by gender, particularly during adolescence.
- **Education engagement** refers to behavioural, emotional and cognitive engagement with schooling. This includes boys' motivational orientations (such as growth mindset, self-efficacy and self-concept), engagement-related behaviours (including homework completion, misbehaviour and help-seeking), and school-related attitudes and feelings (such as fear of failure, enjoyment of school and sense of belonging).

Student-level drivers directly impact **education outcomes** across three interconnected domains.

- **Cognitive outcomes** refer primarily to intellectual abilities. They include foundational skills (literacy, numeracy), domain knowledge and skills (science, civics etc.) and higher-order thinking (such as critical thinking and problem-solving). Where possible, these outcomes are analysed distributionally, with particular attention to low performers, high performers and intersections with other sources of education inequality, such as socio-economic status, language spoken at home and migration status.
- **Non-cognitive outcomes** encompass social-emotional skills and civic attitudes, reflecting the greater availability of international comparative data on these domains. These outcomes interact with cognitive development, shape learning trajectories and societal participation, and are associated with longer-term attainment, wellbeing and labour-market outcomes.
- **Attainment outcomes** capture progression through the education system, completion and achievement (grades and test scores), including indicators of stalled progression such as early school leaving, grade repetition and exclusion. Although grades and standardised test scores technically reflect students' cognitive outcomes, this paper considers them as attainment outcomes as they are often used for determining students' certification and progression through the education system.

The model is intentionally dynamic rather than linear, with context, student-level drivers and educational outcomes linked through feedback loops over time. For example, social-emotional skills shape engagement, behaviour and motivation, while experiences such as disengagement or bullying can in turn undermine wellbeing, allowing early difficulties to compound as students progress through the system.

Rather than treating gender as a stand-alone explanatory factor, the model emphasises converging risk profiles, where particular combinations of contextual conditions and student-level drivers interact to increase the likelihood of poor outcomes. This perspective helps explain why gender gaps are often modest on average but more pronounced among low-performing students, and why boys are disproportionately represented among those facing the most severe educational risks.

Data sources

This literature review uses peer-reviewed studies, national data analysis and international large-scale assessments and surveys (ILSAS - Table 1). ILSAS form the backbone of Part I. This is because international large-scale assessments such as PISA, PIRLS and TIMSS are the primary source of comparative evidence on gender gaps in education. They apply standardised instruments and background questionnaires to representative national samples, enabling meaningful comparisons across countries and over time. Unlike national assessments, which use different frameworks, grading scales and content, these surveys allow researchers to identify patterns that are genuinely cross-national rather than artefacts of particular assessment systems. Although ILSAS are the gold standard in terms of allowing comparison of vastly different education systems across the world, their use comes with some caveats. What follows is a non-exhaustive list.

Table 1 International large-scale assessments and surveys used in this literature review

Organisation	Description	Rounds	Acronym
International Association for the Evaluation of Educational Achievement (IEA)	The latest round of the Trends in Mathematics and Science Study (TIMSS) was conducted in 2023 across 64 countries and 6 benchmarking systems. TIMSS has two target populations: the fourth grade (average age 9.5 years) and the eighth grade (average age 13.5 years). At fourth grade, nearly 360,000 students took part. At eighth grade, almost 300,000 students took part (Von Davier et al., 2024).	1995, 1999, 2003, 2007, 2011, 2015, 2019, 2023.	TIMSS
IEA	The latest round of the Progress in International Reading Literacy Study (PIRLS) was carried out in 2021 across 57 countries, gathering data from about 400,000 students aged around 10, 380,000 parents, 20,000 teachers, and 13,000 schools (Mullis et al., 2023). PIRLS tests students in fourth grade (average age 9.5 years).	2001, 2006, 2011, 2016, 2021.	PIRLS
IEA	The International Civic and Citizenship Education Study (ICCS) reports on students' knowledge and understanding of concepts and issues related to civics and citizenship, as well as their beliefs, attitudes, and behaviours respect to this domain. ICCS 2022 was the latest round, involving 3,400 School Principals, 40,000 Teachers and 82,000 Eighth Grade (aged 13/14) Students across 24 countries (IEA & ACER, 2024).	2009, 2016, 2022.	ICCS

Organisation	Description	Rounds	Acronym
IEA	The International Computer and Information Literacy Study (ICILS) contains a core assessment of student's computer and information literacy skills. Countries also have an option for their students to complete an assessment of computational thinking. The 2023 study assessed over 130,000 students at eighth grade (average age 13.5 years) and gathered data from around 5,000 schools and 60,000 teachers in 35 countries (IEA, 2025).	2013, 2018, 2023.	ICILS
Organisation for Economic Cooperation and Development (OECD)	PISA students are 15-years-old. For the latest round of PISA in 2022, eighty-one countries and economies took part and tested nearly 690,000 students. The assessment which primarily covers reading, mathematics and science, although additional components, such as problem solving, global competence, creative thinking, financial literacy are available for specific rounds of testing (OECD, 2023b).	2000, 2003, 2006, 2009, 2012, 2015, 2018, 2022.	
OECD	The Survey on Social and Emotional Skills (SSES) is the largest international survey that analyses the development of social and emotional skills for 10- and 15-year-old students. In the 2023 round, 16 Countries/Cities implemented the collected contextual information from approximately 18,000 10-year-olds and 52,000 15-year-olds, as well as teachers, school principals, and parents (OECD, 2025c).	2019, 2023.	SSES

Notes: Countries that participated in the individual rounds of each study can be found on the IEA and OECD websites at <https://www.iea.nl/studies> and <https://www.oecd.org/en/about/members-partners.html>

Firstly, since ILSAS are cross-sectional, they only provide a snapshot of performance in that specific assessment year. They are, by and large, not designed to track the same students over time and the results are correlational, not causal. As such, they are generally unable to diagnose which specific aspects of education policy or practice affect the trajectories of individual students. To mitigate this, quasi-experimental designs such as randomised control trials, are used alongside the more qualitative and descriptive studies in parts II and III of the literature review.

Secondly, ILSAS tend to be regarded as low stakes⁴ in that their test scores have no impact on students' progression through the education system and/or grades. There are differences in terms

⁴ There are country differences in how the stakes are perceived. For example, in Asian and North European countries students tend to invest considerable effort to perform well in PISA (Griselda, 2024).

of how the same students perform in high and low-stakes exams, which tend to be larger for boys, Caucasians and higher SES students (Akyol et al., 2021). Analysis of PIRLS data from Finland, Germany, Singapore, and the United Arab Emirates found boys tended to be less engaged and faster at completing the test, compared to girls but those boys who were more engaged and took their time had similar results to girls (Leng & Von Davier, 2025). There is also consistent evidence that boys tend to skip more questions in low-stakes tests compared to girls, due to lower motivation and effort in low-stakes environments (Eklöf, 2007; Gafni & Melamed, 1994; Marcq & Braeken, 2024). To mitigate this, patterns seen in ILSAS are analysed in addition to national grades and test scores in high-stakes conditions.

Thirdly, student assessments use a random assignment of different assessment booklets, which are all designed to assess the same level of proficiency but have different proportions of question types. In PISA, as little as 20% of the questions in some booklets are multiple-choice, in others they can be as high as 70% of the test. Research has found that gender gaps in favour of girls are larger for open-ended questions (Beller & Gafni, 2000). Analysis of PISA data shows that a 10 percentage point increase in the share of multiple-choice questions inflates female under-performance in mathematics and male under-performance in reading by nearly one-quarter of the overall gender disparity found in these domains (Griselda, 2024). Previous research has suggested that the gender gap in favour of girls is larger for assessments that consider print reading vs digital reading (Borgonovi, 2016). However, subsequent studies have found that advantages in digital reading are largely reflective of differences in overall reading skills, rather than a digital-specific advantage for one gender or the other (Grammatikopoulou et al., 2026; Mirazchiyski & Gershteyn, 2024).

Another important caveat of ILSAS is the use of exclusions by participating countries. In addition to excluding home-schooled or permanently suspended students, countries are allowed to exclude entire schools or individual students with special education needs or additional language requirements. Importantly, students cannot be excluded solely because of low proficiency or common disciplinary problems, and the most common reason is intellectual disability. When it comes to PISA, assessment standards dictate that the overall exclusion rate within a country/economy stays below 5% of the target sample (OECD, 2024). However, in the latest round of PISA, sixteen countries/economies failed to meet this standard. Among those with the highest exclusion rates were Denmark (11.6%), the Netherlands (8.4%), Latvia (7.9%), Sweden (7.4%), and Norway (7.3%). In the case of Denmark, exclusions showed a marked increase compared to 2018 (5.7%), a major cause of this increase was the increased share of students with diagnosed dyslexia (OECD, 2023b).

There are important implications to these sampling differences. In Sweden, research found that the PISA sample in 2022 overestimated the actual achievement level by disproportionately excluding students with weaker academic performance (Borger et al., 2024). While for PIRLS, TIMSS and ICCS, there is a modest association between exclusion rates and achievement at the country level, additional data is needed to assess the extent of any bias (Atasever et al., 2024). International comparisons should therefore be made cautiously, since some countries (such as Japan and Korea) meet the sampling criteria to a much greater extent than others. Even in countries that meet the standards more rigidly, studies have also raised concerns of upward bias in PISA results. For example, in Canadian and British PISA samples (Anders et al., 2021; Durrant & Schnepf, 2018; Jerrim, 2021; Micklewright et al., 2012).

Despite these valid caveats, ILSAS still represent the only internationally comparative datasets of cognitive skills, drawing on millions of test takers in dozens of countries all over the world.

References

Akyol, P., Krishna, K., & Wang, J. (2021). Taking PISA Seriously: How Accurate are Low-Stakes Exams? *Journal of Labor Research*, 42(2), 184–243. <https://doi.org/10.1007/s12122-021-09317-8>

Anders, J., Has, S., Jerrim, J., Shure, N., & Zieger, L. (2021). Is Canada really an education superpower? The impact of non-participation on results from PISA 2015. *Educational Assessment, Evaluation and Accountability*, 33(1), 229–249. <https://doi.org/10.1007/s11092-020-09329-5>

Arnold, R. D., & Wade, J. P. (2015). A Definition of Systems Thinking: A Systems Approach. *Procedia Computer Science*, 44, 669–678. <https://doi.org/10.1016/j.procs.2015.03.050>

Atasever, U., Jerrim, J., & Tieck, S. (2024). Exclusion rates from international large-scale assessments: An analysis of 20 years of IEA data. *Educational Assessment, Evaluation and Accountability*, 36(3), 405–428. <https://doi.org/10.1007/s11092-023-09416-3>

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2023). Males at the Tails: How Socioeconomic Status Shapes the Gender Gap. *The Economic Journal*, 133(656), 3136–3152. <https://doi.org/10.1093/ej/uead069>

Beller, M., & Gafni, N. (2000). Can Item Format (Multiple Choice vs. Open-Ended) Account for Gender Differences in Mathematics Achievement? *Sex Roles*, 42(1–2), 1–21. <https://doi.org/10.1023/A:1007051109754>

Borger, L., Johansson, S., & Strietholt, R. (2024). How representative is the Swedish PISA sample? A comparison of PISA and register data. *Educational Assessment, Evaluation and Accountability*, 36(3), 365–383. <https://doi.org/10.1007/s11092-024-09438-5>

Borgonovi, F. (2016). Video gaming and gender differences in digital and printed reading performance among 15-year-olds students in 26 countries. *Journal of Adolescence*, 48(1), 45–61. <https://doi.org/10.1016/j.adolescence.2016.01.004>

Bronfenbrenner, U. (1981). *The Ecology of Human Development: Experiments by Nature and Design*. Harvard University Press. <https://doi.org/10.2307/j.ctv26071r6>

Brozo, W. G. (2019). *Engaging boys in active literacy: Evidence and practice*. Cambridge University Press. <https://doi.org/10.1017/9781108654111>

Durrant, G. B., & Schnepf, S. V. (2018). Which Schools and Pupils Respond to Educational Achievement Surveys?: A Focus on the English Programme for International Student Assessment Sample. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 181(4), 1057–1075. <https://doi.org/10.1111/rssa.12337>

Eklöf, H. (2007). Test-Taking Motivation and Mathematics Performance in TIMSS 2003. *International Journal of Testing*, 7(3), 311–326. <https://doi.org/10.1080/15305050701438074>

Farquharson, C., McNally, S., & Tahir, I. (2024). Education inequalities. *Oxford Open Economics*, 3(Supplement_1), i760–i820. <https://doi.org/10.1093/ooec/odad029>

Gafni, N., & Melamed, E. (1994). Differential tendencies to guess as a function of gender and lingual-cultural reference group. *Studies in Educational Evaluation*, 20(3), 309–319. [https://doi.org/10.1016/0191-491X\(94\)90018-3](https://doi.org/10.1016/0191-491X(94)90018-3)

Gilbert, L., Teravainen, A., Clark, C., & Shaw, S. (2018). *Literacy and life expectancy*. A National Literacy Trust research report. https://nlt.hacdn.org/media/documents/National_Literacy_Trust_-_Literacy_and_life_expectancy_report.pdf

Grammatikopoulou, E., Johansson, S., & Rosén, M. (2026). Paper-based and digital reading in 14 countries: Exploring cross-country variation in mode effects. *Educational Review*, 78(1), 61–79. <https://doi.org/10.1080/00131911.2025.2452236>

Gregory, L., & Welmond, M. J. (2021). *Educational Underachievement Among Boys and Men (English)*. World Bank Group. <http://documents.worldbank.org/curated/en/111041644611110155>

Griselda, S. (2024). Gender gap in standardized tests: What are we measuring? *Journal of Economic Behavior & Organization*, 221, 191–229. <https://doi.org/10.1016/j.jebo.2024.03.010>

Hammond, A., Rubiano Matulevich, E., Beegle, K., & Kumaraswamy, S. K. (2020). *The Equality Equation*. World Bank, Washington, DC. <https://doi.org/10.1596/34317>

Hattie, J. (2008). *Visible Learning* (0 edn). Routledge. <https://doi.org/10.4324/9780203887332>

Hattie, J., & Anderman, E. M. (Eds). (2019). Gender Influences. In J. Gill, *Visible Learning Guide to Student Achievement* (1st edn, pp. 6–11). Routledge. <https://doi.org/10.4324/9781351257848-2>

Hyde, J. S. (2014). Gender Similarities and Differences. *Annual Review of Psychology*, 65(1), 373–398. <https://doi.org/10.1146/annurev-psych-010213-115057>

IEA. (2025). ICILS 2023 International Database: International Computer and Information Literacy Study_Edition_2 (Version 2) [Data set]. IEA - International Association for the Evaluation of Educational Achievement. https://doi.org/10.58150/ICILS_2023_DATA_2

IEA & ACER. (2024). *ICCS 2022 International Database: International Civic and Citizenship Education Study - Edition 2* [Data set]. IEA. https://doi.org/10.58150/ICCS_2022_DATA_EDITION_2_INCLUDING_PROCESS_DATA

Jerrim, J. (2021). PISA 2018 in England, Northern Ireland, Scotland and Wales: Is the data really representative of all four corners of the UK? *Review of Education*, 9(3), e3270. <https://doi.org/10.1002/rev3.3270>

Jha, J., Menon, N., & Chatterjee, D. (2017). *Boys' Underperformance in the Commonwealth: Revisiting the Issue*. Commonwealth of Learning. <https://oasis.col.org/server/api/core/bitstreams/0654d268-ee6f-42dc-a279-c888643018fa/content>

Leng, D., & Von Davier, M. (2025). Gender Differences and Types of Test-Taking Behaviors in PIRLS 2021. *Applied Measurement in Education*, 38(1), 75–93. <https://doi.org/10.1080/08957347.2025.2486032>

Marcq, K., & Braeken, J. (2024). Gender differences in item nonresponse in the PISA 2018 student questionnaire. *Educational Assessment, Evaluation and Accountability*, 36(3), 301–336. <https://doi.org/10.1007/s11092-023-09412-7>

Micklewright, J., Schnepf, S. V., & Skinner, C. (2012). Non-Response Biases in Surveys of Schoolchildren: The Case of the English Programme for International Student Assessment (PISA) samples. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 175(4), 915–938. <https://doi.org/10.1111/j.1467-985X.2012.01036.x>

- Mirazchiyski, P. V., & Gershteyn, V. (2024). Technology factors related to the differences in paper and online reading scores in PIRLS 2016. *Large-Scale Assessments in Education*, 12(1), 33. <https://doi.org/10.1186/s40536-024-00224-9>
- Mullis, I., Von Davier, M., Foy, P., Fishbein, B., Reynolds, K., & Wry, E. (2023). *PIRLS 2021 International Results in Reading*. TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.tr2103.kb5342>
- OECD. (2023a). *PISA 2022 Assessment and Analytical Framework*. OECD Publishing. <https://doi.org/10.1787/dfe0bf9c-en>
- OECD. (2023b). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- OECD. (2024). *PISA 2022 Technical Report*. OECD. <https://doi.org/10.1787/01820d6d-en>
- OECD. (2025a). *Gender Equality in a Changing World: Taking Stock and Moving Forward*. OECD Publishing. <https://doi.org/10.1787/e808086f-en>
- OECD. (2025b). *Gender wage gap* [Data set]. <https://www.oecd.org/en/data/indicators/gender-wage-gap.html>
- OECD. (2025c). *OECD Survey on Social and Emotional Skills 2023 Technical Report*. OECD Publishing. <https://doi.org/10.1787/c37f1703-en>
- Pawson, R., & Tilley, N. (2010). *Realistic evaluation* (Reprinted). SAGE.
- Purves, A. C. (1987). The Evolution of the IEA: A Memoir. *Comparative Education Review*, 31(1), 10–28. <https://doi.org/10.1086/446653>
- Richmond, B. (1994). Systems thinking/system dynamics: Let's just get on with it. *System Dynamics Review*, 10(2–3), 135–157. <https://doi.org/10.1002/sdr.4260100204>
- Roberts, S. (2012). 'I just got on with it': The educational experiences of ordinary, yet overlooked, boys. *British Journal of Sociology of Education*, 33(2), 203–221. <https://doi.org/10.1080/01425692.2011.649832>
- Stenberg, I. (with Boström, L., Bostedt, G., & Edling, S.). (2025). *Bridging the Gender Gap Analysing Boys' School Experiences and Educational Challenges*. Mid Sweden University.
- UNICEF. (2022). *Brief on the Social Ecological Model*. <https://www.unicef.org/media/135011/file/Global%20multisectoral%20operational%20framework.pdf>
- Von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U., & Yin, L. (2024). *TIMSS 2023 International Results in Mathematics and Science*. TIMSS & PIRLS International Study Center, Boston College. <https://doi.org/10.6017/lse.tpisc.timss.rs6460>
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140(4), 1174–1204. <https://doi.org/10.1037/a0036620>
- Younger, M., & Warrington, M. (2005). *Raising boys' achievement* [University of Cambridge Faculty of Education]. DFES Publications.

Published by the Centre for Policy Research on Men and Boys (CPRMB),
May 2026

E: Contact@menandboys.org.uk

W: <https://www.menandboys.org.uk>

LinkedIn: [centre-policy-research-menboys](https://www.linkedin.com/company/centre-policy-research-menboys)

Instagram: [cprmenandboys](https://www.instagram.com/cprmenandboys)

X: [CPRMenBoys](https://twitter.com/CPRMenBoys)

Centre for Policy Research on
Men and Boys

The contents and opinions expressed in this paper are those of the authors only.

Address: Centre for Policy Research on Men and Boys
c/o 2 Dean Trench Street, Westminster, SW1P 3HE

© CPRMB 2026

Publication Number: 02/26